HENP Grand Challenge on Data Access*: Optimizing Retrieval of Events for STAR/RHIC in MDC 1

*D. Olson, J. Yang, D. Zimmerman,† A. Shoshani, L. Bernardo, H. Nordberg, A. Sim,‡ D. Malon,§ R. Porter, T. Wenaus***

The HENP Grand Challenge on Data Access project has been developing techniques and software for optimizing the retrieval of event data from a robotic tape storage system (HPSS). The first use of this system was carried out during the RHIC Mock Data Challenge 1 at Brookhaven National Laboratory in Sept./Oct. 1998.

The software is implemented with a set of CORBA servers, called the storage manager, that accept queries from the clients (physicists' analysis programs) and retrieve the necessary files from HPSS. A query specification consists of a selection string, e.g., "num_pion>500" or "mean_pt<4.&num_p>50", or a list of events. The storage manager determines the set of files required for each query and which events in each file qualify. The order that files are retrieved from the HPSS system is determined by a set of policy parameters that can be tuned for optimum performance. In the example shown in figure 1 files are retrieved in order of number of events to be used in each file as well as the number of queries that need access to a file. Events are returned to the client analysis program via an order optimized iterator as each file is staged to disk. From the point of view of the analysis program, there is a single loop over all events that satisfy the query conditions while the files are being transparently retrieved from tape to disk.

This system is enhanced to include event components distributed across multiple files for RHIC MDC2 in March 1999.

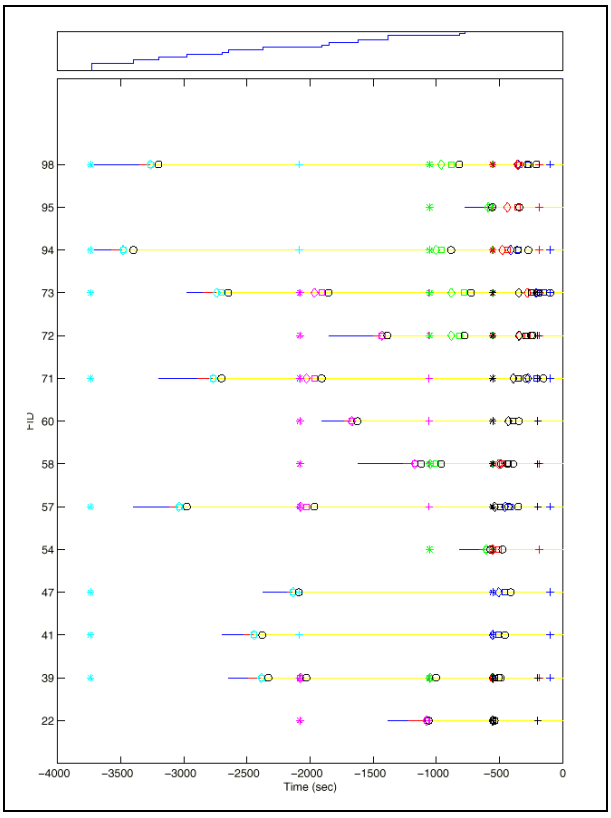

Fig. 1. Illustrates retrieving DST event files for STAR from the RCF HPSS system during the Mock Data Challenge in Sept. 1998. Three queries execute with a staggered start time (cyan, magenta, green) and repeated with a simultaneous start (blue, black, red). The time, in seconds before the finish, is indicated on the horizontal axis. The file number is indicated on the vertical axis. A blue horizontal line indicates the time between a request for a file and time that the data transfer across the network begins. The red line indicates the duration of network transfer. The yellow line indicates that a file resides on disk, to be shared by queries.
The blue histogram at the top indicates the amount of space used in the disk cache.

Footnotes and References

* http://www-rnc.lbl.gov/GC/

† Nuclear Science Division, LBNL

‡ NERSC, LBNL

§ Argonne National Laboratory

** Brookhaven National Laboratory